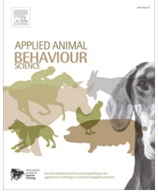
Applied Animal Behaviour Science 232 (2020) 105099

Contents lists available at ScienceDirect

Applied Animal Behaviour Science

journal homepage: www.elsevier.com/locate/applanim



Stereotypic behaviour in horses lowers stress but not spatial learning performance

S. Briefer Freymond[a,*], S. Beuret[b], A. Ruet[a], K. Zuberbühler[b,c], I. Bachmann[a], E.F. Briefer[d]

[a] Swiss National Stud Farm, Les Longs-Prés, 1580, Avenches, Switzerland
[b] Institute of Biology, University of Neuchâtel, rue Emile-Argand 11, 2000, Neuchâtel, Switzerland
[c] School of Psychology & Neurosciences, University of St Andrews, Scotland, United Kingdom
[d] Behavioural Ecology Group, Section for Ecology & Evolution, Department of Biology, University of Copenhagen, 2100, Copenhagen Ø, Denmark

ARTICLE INFO

Keywords:
Crib-biting
Cortisol
Learning
Frustration

ABSTRACT

Stereotypies are common in captive animals, but it remains unclear if they are pathological by-products of captive conditions or if they have an adaptive function. Here, we address this question using crib-biting, a common type of stereotypic behaviour in domestic horses, thought to result from stressful or frustrating environments. Since chronic stress is known to affect learning abilities via dopaminergic modulation in the basal ganglia, we predicted that stereotypic horses would underperform in learning tasks compared to healthy control animals, in line with the pathological by-product hypothesis. We exposed 19 crib-biters and 18 control horses in six spatial tasks, and collected behavioural and physiological data (heart rate, heart rate variability and salivary cortisol) to assess learning performance and stress levels. We found that, during the learning phase, 10 of 19 crib-biters showed stereotypic behaviour. Surprisingly, crib-biters that displayed the stereotypic behaviour (Group A) had lower salivary cortisol values compared to crib-biters that did not display the stereotypic behaviour (Group B) or control horses (Group C), after the first part of the experiment. In addition, a smaller proportion of horses in Group A displayed behaviours indicating frustration during one of the test compared to Group B. Moreover, we found no difference between the three groups in their learning performance nor in their heart rate or heart rate variability measures. Our results therefore suggest that crib-biting has an adaptive function and can help horses to reduce stress and frustration during learning tasks, which then enables them to show normal learning performance.

1. Introduction

Stereotypies in animals have been defined as repetitive, relatively invariant patterns of behaviour, with no apparent goal or function (Mason and Latham, 2004). Stereotypies also exist in humans where they can be either psychologically or environmentally induced and are often associated with developmental disorders, such as autism, neurological disorders, obsessive compulsive disorder, Tourette's syndrome and severe psychiatric disturbances (e.g. schizophrenia) (McBride and Parker, 2015). In animals, they can appear as fixed locomotor patterns (e.g. compulsive digging and looping in gerbils and mice, *Mus musculus* (Würbel and Stauffacher, 1994)) or fixed oral movements (e.g. bar-biting and sham-chewing in sows, *Sus scrofa*, tongue rolling in cows, *Bos Taurus* (Sambraus, 1985)). Animal stereotypies appear to be mainly induced by restricted environments, suboptimal housing conditions or management problems (McBride and Parker, 2015).

Stereotypies have been reported to result from stressful or frustrating environments (Luescher et al., 1991; McBride and Parker, 2015). Internal or external stressors normally induce necessary and adaptive physiological reactions (von Borell, 2001). These reactions help individuals to cope with a situation and regain homeostasis (or a steady state). However, stressors can differ qualitatively (physical or psychological) and quantitatively (chronic, acute, intermittent) (McBride and Parker, 2015). When stressors, such as unpredictable, repetitive

---

Abbreviations: LMMs, Linear mixed-effects models; GLMMs, Generalised linear mixed models; HR, Heart rate; RMSSD, Root mean square of successive inter-beat interval differences; ECG, Electrocardiogram; HPA, Hypothalamo-pituitary adrenocortical axis; SAM, Neural sympatho-adreno-medullary axis; ACTH, Adenocorticotropin hormone; RR, Inter-heartbeat intervals.

[*] Corresponding author at: Swiss National Stud Farm SNSF, Les Longs Prés, CH-1580, Avenches, Switzerland.
E-mail address: sabrina.briefer@agroscope.admin.ch (S. Briefer Freymond).

https://doi.org/10.1016/j.applanim.2020.105099
Received 6 January 2020; Received in revised form 16 July 2020; Accepted 25 August 2020
Available online 28 September 2020
0168-1591/

situations, are sustained for an extended period, a chronic stress state can occur (Koolhaas, 2011), which can have deleterious effects on animals, including the development of stereotypies. Physiologically, two pathways are activated during a stress response; the hormonal hypothalamo-pituitary-adrenocortical (HPA) axis and the neural sympatho-adreno-medullary (SAM) axis. Chronically stressed animals can develop, for example, an increased or maintained HPA responses to any novel stressor (Bhatnagar and Vining, 2003).

Following long-term chronic stress exposure, stress hormones can affect the neurobiological system of stereotypic animals (Cabib and Puglisi-Allegra, 2012). More specifically, the functioning of the *basal ganglia* can be modified through alterations of the dopaminergic system (Saka et al., 2004; McBride and Hemmings, 2005). This has been notably described for crib-biter horses, which have been reported to have structural alterations in the *striatum* of the *basal ganglia*, with higher numbers of dopamine receptors in the *ventral striatum* and lower numbers in the *dorsomedial striatum*, resulting in sensitisation to dopamine release (McBride and Hemmings, 2005). Similar neurophysiological differences within the *basal ganglia* have been identified in stereotypic mice (Cabib et al., 1998; McBride and Parker, 2015). Dopaminergic dysfunction is, in turn, thought to have cognitive implications. For example, an excessive desire for food rewards might occur, which will have an impact on cognition tests or trigger a lack of flexibility (e.g. perseveration) impacting on learning performances and predispositions to frustration (Ijichi et al., 2013). In addition, short-term acute stress can have a negative impact on learning performances and memory (Valenchon et al., 2013), mediated by HPA-released glucocorticoids and SAM-released catecholamines (Schwabe and Wolf, 2010).

Support for a role of *basal ganglia* dysfunction in learning has been shown in autistic people, which show poor abilities to suppress behaviours (perseveration), although high stereotypic levels do not always correlate with higher perseveration (Lopez et al., 2005; Boyd et al., 2009). In schizophrenic and autist patients, a correlation between inappropriately repeated responses, “recurrent perseveration”, and the intensity of the stereotypy was shown in two-choice guessing tasks (Frith and Done, 1983; Turner, 1997). In animals, stereotypy levels (i.e., the frequency of stereotypic performance) are also strong predictors of the latency to extinguish conditioned responses, or of the tendency to inappropriately repeat responses, which are both considered as indirect measures of *basal ganglia* dysfunction in several species, such as bears, *Ursus thibetanus* and *Helactos malayanus* (Vickery and Mason, 2005), orange-wing amazon parrots, *Amazona amazonica* (Garner et al., 2003) and bank voles, *Clethrionomys glareolus* (Garner and Mason, 2002). This further suggests a relationship between dopaminergic dysfunction and lack of cognitive flexibility (Garner and Mason, 2002; McBride and Parker, 2015). Since the extinction of a learned response that no longer leads to a desired outcome is a cognitive process that is essential in response to a changing environment, this implies that stereotypies should be considered maladaptative (Cooper and Mason, 1998; Valenchon et al., 2013).

In domestic horses, oral (crib-biting or windsucking) and locomotor stereotypies (weaving and box-walking) are the most common repetitive abnormal behaviours (Mason and Rushen, 2006). For instance, crib-biting has a prevalence of 2.4–8.4 % (Giuliana and Roberto, 1986; Luescher et al., 1991; Albright et al., 2009; Muñoz et al., 2018). Similar to stereotypic bears, crib-biting horses persist longer during the extinction phase of operant learning tasks than controls (Hemmings et al., 2007; Roberts et al., 2015). Moreover, crib-biters were suggested to be more prone to habit formation after overtraining (Parker et al., 2008, 2009). Overall, the current consensus is that stereotypies may favour habit formation and thus a loss in cognitive flexibility (Schwabe and Wolf, 2011). On the other hand, a recent study failed to show evidence of learning impairment between crib-biters and control horses in a visual reversal learning task, which has been considered as an indirect measure of flexibility of learning and attention (Briefer Freymond et al., 2018). However, most of these results are based on limited sample sizes (range 4–10 horses). In addition, locomotor detour tasks, which are generally used as measures of impulsivity or behavioural/motor inhibition (Jentsch et al., 2000; Kabadayi et al., 2017), have to our knowledge never been applied in stereotypic horses. In such tasks, animals need to inhibit the propensity of directly reaching a reward and instead have to make a detour (Dallaire et al., 2011; Kabadayi et al., 2017).

In this study, we compared the performance of a relatively large number ($N = 19$) of horses displaying stereotypic behaviour (crib-biters) and a similar number of non-stereotypic control individuals ($N = 18$) in frustration-causing spatial learning tasks, i.e., reversal, extinction, and detour learning. Frustration in learning tasks can occur when previously learned expectations are no longer met, or when sudden barriers prevent access to a reward (Psyrdellis et al., 2016). Based on previous results, we predicted that crib-biters would show a stronger dependency on rewards and respond differently to the tests due to alterations in the dopaminergic system in some parts of the *basal ganglia* (McBride and Hemmings, 2005; Hemmings et al., 2007; Parker et al., 2008, 2009; Ijichi et al., 2013; Roberts et al., 2015; Hemmings et al., 2018). Additionally, according to previous results (Briefer Freymond et al., 2015), we expected crib-biters to show higher physiological stress responses during the experiment than controls, but that executing stereotypic behaviours would have a beneficial effect on reducing stress and also potentially on learning capacities.

2. Methods

2.1. Subjects and management conditions

The study was carried out from September 2013 to February 2014 on 19 crib-biters and 18 control horses of various breeds, sexes (mares, geldings and stallions) and ages (3–24 years old; Table 1), housed in 19 different farms in Switzerland (Table 1). Except for one control, all horses participated in a previous study aimed at testing the physiological reaction of crib-biters and non-stereotypic horses in a standard ACTH challenge test (Briefer Freymond et al., 2015). Twenty-six horses were privately owned and 11 belonged to the Swiss National Stud Farm (https://www.agroscope.admin.ch/agroscope/en/home/about-us/snsf.html). All subjects had lived on their premises for at least one year prior to testing. To qualify as a crib-biter, a subject had to have shown crib-biting behaviour for a minimum of one year, according to its owner. Controls were horses that had never been observed crib-biting or showing other stereotypies (i.e. weaving, box walking, head tossing nodding (Lesimple and Hausberger, 2014)). This grouping was verified in a first study (Briefer Freymond et al., 2015) and further confirmed during this study (i.e. crib-biters were all observed crib biting at some point before or during our tests, while control horses were never seen displaying any stereotypy). We matched each crib-biting horse with a control horse of similar breed, sex and age, housed in similar conditions and if possible on the same farm (Table 1). Housing was either as individuals or in groups, in single boxes or paddocks, depending on the farm (Table 1). Routine care was provided by the owners. The study was approved by the Swiss Federal Veterinary Office (approval number VD 2677 bis; Switzerland). The owners of the horses were provided with a detailed written description of the experiment to be conducted and agreed to the research being carried out on their animals.

2.2. Experimental procedure

The first part of this experiment, aimed at characterising the personality of crib-biters, is described in another paper (Briefer Freymond et al., 2019).

During each test day, one horse (9 farms) or two horses (7 farms) per farm (one crib-biter and one control; Table 1) were successively tested at around 11am. Before starting the tests, a wireless heart-rate monitor was fixed to a surcingle strapped around the subject’s heart girth, and a saliva sample was taken in the home pen. The experimental procedure,

Table 1
Characteristics of the horses used in the experiment. Sex (M = mare; G = gelding, S = stallion), group (CB = crib-biters; C = controls; A = crib-biters that did crib-bite; B = crib-biters that did not crib-bite), part 1 and 2 of the learning tests, age, breed (FM = Franches-Montagnes; ET = English Thoroughbred), and farm (each letter refers to a given farm). Group ABC indicates which sub-group the horses were part of (Group A corresponds to the crib-biters that crib-bit during part 1 and/or during part 2 of the learning test, Group B to the crib-biters that did not crib-bite during part 1 and/or during part 2 of the learning test, and Group C to the controls) during part 1 and part 2 of the experiment.

Horses	Sex	Group CB-C	Group ABC part 1 /part 2	Age (years)	Breed	Place
1	M	CB	B - B	6	Swiss Warmblood	c
2	M	CB	B - A	22	Criollo	g
3	M	CB	A - A	16	FM	y
4	M	CB	A – A	9	Hispano-Arabian	b
5	M	CB	B – B	5	Quarter Horse	s
6	M	CB	A – A	9	Paint Horse	r
7	M	CB	A – A	5	Paint Horse	k
8	G	CB	A – B	9	FM	d
9	G	CB	A – B	11	Swiss Warmblood	g
10	G	CB	B – A	23	FM	n
11	G	CB	A – A	11	FM	bo
12	S	CB	B – B	17	FM	h
13	S	CB	B – B	15	FM	h
14	M	CB	A – A	5	FM	m
15	G	CB	B – A	19	Haflinger	se
16	G	CB	A – A	18	Swiss Warmblood	a
17	G	CB	B – A	7	Unknown origin	v
18	G	CB	A – A	10	ET	d
19	S	CB	B – B	11	FM	h
20	M	C	C	7	Quarter Horse	s
21	M	C	C	20	FM	y
22	M	C	C	14	Swiss Warmblood	h
23	M	C	C	18	Camargue	b
24	M	C	C	14	Swiss Warmblood	h
25	M	C	C	16	Trotter	h
26	M	C	C	18	FM	h
27	M	C	C	10	Swiss Warmblood	g
28	G	C	C	4	Friso-Arabian	n
29	G	C	C	24	Unknown origin	v
30	G	C	C	22	ET	d
31	G	C	C	7	Quarter Horse	k
32	G	C	C	6	FM	di
33	G	C	C	8	FM	d
34	G	C	C	15	Swiss Warmblood	h
35	G	C	C	11	Swiss Warmblood	h
36	S	C	C	17	FM	h
37	S	C	C	7	FM	h

based on preliminary tests performed with 20 pilot horses (different horses as those used in this study), was as follows: the subject was led from its home pen to an 8 × 10 m familiar test arena for a 15 min habituation period to the experimental material before being subjected to six learning tests (Figs. 1 and 2). A daily session began with three spatial learning tests (part 1, Fig. 2), performed successively over a period of 60–70 min (along with a habituation phase and personality tests that are described in another paper (Briefer Freymond et al., 2019)). After this first part, the subject was returned to its home pen for a second saliva sample and a pause of 60–70 min. At the end of the pause, the three remaining spatial learning tests (part 2, Fig. 2) were performed for a duration of 10–15 min before it was led back to its home pen for a third saliva sample. Tests were always conducted in the same order. At the beginning of each trial, experimenter 1 led the subject to the start, said "top", and released the horse from its lead. All the tests were video recorded with a camera (Sony Handycam HDR-CX700) by a second experimenter situated outside the arena.

2.2.1. *Test 1: place learning*

Test 1 was performed in order to teach the horses to associate a food reward (horse commercial concentrate) with a bucket and as requirement for the next test (Test 2). It consisted of a fixed number of eight trials, during which a subject could learn that the food was always situated in a bucket placed on one side of the arena. To this end, a solid fence was placed in the middle of the arena in order to divide it into two equal parts. Then, two identical buckets, one empty and one filled with grain (*Hypona 788–2*), were randomly assigned across subjects to each side of the fence (Fig. 2). Experimenter 1 positioned the horse in front of the fence before releasing it alternatively either to her left or right side. The horse was then allowed to explore the arena for up to three minutes to find the baited bucket. If it did so within a maximum of three minutes, it was allowed to take a few bite of grain; otherwise it was returned to the starting position by the experimenter for the next trial without eating. During each trial, experimenter 2 recorded through direct observation the time required to find the baited bucket, the side of the fence selected by the horse after leaving the start zone, and whether or not it inspected the empty bucket. We predicted that both groups would quickly learn, as shown by a decrease in the time needed to go to the baited bucket (Parker et al., 2009).

2.2.2. *Test 2: reversal learning*

Test 2 was performed to obtain a general measure of cognitive flexibility for both groups (crib-biters and controls) (Izquierdo et al., 2017; McBride et al., 2017). It consisted of a fixed number of eight trials, during which the horses could learn that the food was no longer available in the bucket situated on the previously assigned side, but instead on the opposite side (Fig. 2). The experimental procedure and data collection were the same as in Test 1. Crib-biters, due to their potential propensity for habit formation, were expected to perform poorly in this test compared to non-stereotypic horses (Hemmings et al., 2007; Parker et al., 2008, 2009; Ijichi et al., 2013; Roberts et al., 2015). Therefore, we expected them to go more often to investigate the empty bucket.

2.2.3. *Test 3: detour learning*

In order to compare the tolerance to frustration and stress sensitivity of crib-biters and control horses, the horses were tested in a detour task. This third test, also called *continuously visible goal* detours, consisted of a fixed number of five trials, during which a subject could learn that it had to move around a fence to reach a bucket with food (Wolff and Hausberger, 1996; Kabadayi et al., 2017). The bucket was placed at one metre from the fence, on the opposite side from which the horse was released, which was randomly assigned across subjects (Fig. 2). Experimenter 1 positioned the horse at one metre perpendicular to the fence before releasing it alternatively from her left or right side. The horse was then allowed to explore the arena for up to three minutes. If it reached the bucket within this timeframe, it was allowed to take a few bite of grain before being returned to the start position for the next trial; otherwise it was directly returned to the starting position without eating. Experimenter 2 scored the time required to reach the bucket through direct observation. We predicted that crib-biters would have more difficulties in performing this detour task compared to non-stereotypic horses, due to their lower tolerance to frustration and increased stress sensitivity. We thus predicted that they would take more time on average to go around the fence in order to reach the baited bucket (Ijichi et al., 2013; Briefer Freymond et al., 2015).

2.2.4. *Test 4: food motivation*

Test 4 was performed in order to assess if one group (crib-biters or

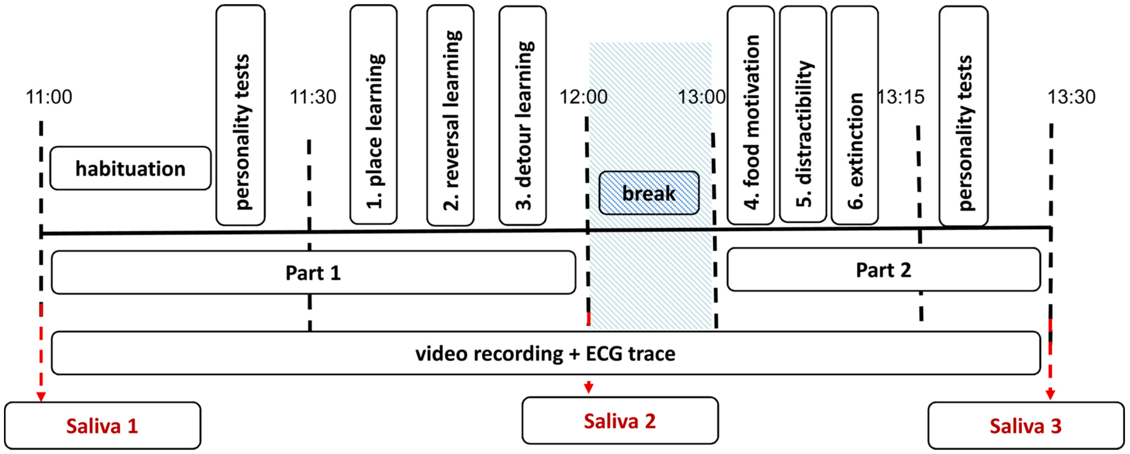


Fig. 1. Experimental procedure for the learning tests. The black dotted lines indicate the time at which each period started and ended. The learning tests are indicated with number (1–6). "Saliva" indicates when the saliva samples for the cortisol measurements were collected (1–3). Video and ECG traces were recorded continuously. Learning tests were performed between two sets of personality tests, which are described in another paper (Briefer Freymond et al., 2019).

controls) was more motivated to eat than the other group. This test consisted of a fixed number of three trials, during which subjects had to learn that one bucket containing food was situated at the end of a corridor (established by a fence placed in the arena; Fig. 2). After being released by experimenter 1 (alternatively from her left or right side), the horse was allowed to explore the arena for three minutes. If it reached the bucket within this timeframe, it was allowed to take a few bite of grain before being returned to the start position for the next trial; otherwise it was directly returned to the starting position without eating. Across trials, we scored the time to reach the bucket through direct observation. We predicted that the food motivation of both group would be high and that they would thus learn quickly, resulting in a short time to go to the baited bucket (Parker et al., 2009; Ijichi et al., 2013).

2.2.5. *Test 5: distractibility*

Test 5 was developed in order to compare the distractibility of crib-biters and control horses. This test was a continuation of Test 4 and consisted of one single trial, during which subjects had to reach the baited food bucket. However, this time, a piece of wood was placed on one side of the corridor (Fig. 2). After being released, the subject was allowed 30 s to explore the arena. Experimenter 2 recorded through direct observation the time the horse took to reach the bucket and if it looked at or touched the piece of wood, as a measure of distractibility. We predicted that crib-biters would show limited exploration of the object, i.e. devote less attention to the piece of wood than control individuals, as shown in Parker et al. (2009) and in Ijichi et al. (2013).

2.2.6. *Test 6: extinction*

Test 6 consisted in an extinction task, aimed at comparing the perseverance of crib-biters and controls. This test consisted of a fixed number of eight trials, during which subjects could learn that there was no more food in the bucket situated at the end of the corridor. After being released alternatively from experimenter 1′s left or right side, subjects were allowed to explore the arena for 30 s, before being returned to the start for the next trial. Experimenter 2 recorded through direct observations the time during which subjects waited at the entrance of the corridor, and whether or not subjects went to look into the bucket. We predicted that crib-biters would show more perseverance by visiting the empty bucket more often and for longer than controls, due to their potentially stonger propensity to form habits and their stronger dependency on rewards, and would thus show less propensity to wait at the entrance of the corridor (Hemmings et al., 2007; Ijichi et al., 2013; Roberts et al., 2015).

2.3. *Response measures*

2.3.1. *Behaviours (scored through direct observation)*

In addition to the behaviours scored through direct observation and described above for each test (summarised in Table 3), experimenter 2 noted on the spot all occurrences of stereotypic behaviour seen during the tests (crib-biting, windsucking) (Table 2). Crib-biting occurrences were then verified by an independent observer, blind to the groups, based on the videos of all the horses. Crib-biting was defined as a subject grasping the solid fence with its incisors, pulling back, contracting the neck muscles and drawing air into its oesophagus with an audible grunt (McGreevy et al., 1995). Windsucking was defined in the same way, but without grasping an object (McGreevy et al., 1995).

2.3.2. *Behaviours (scored from the videos)*

The tests that could potentially induce frustration (*Test 1, 2, 3 and 6*), due to the fact that previously learned expectations are no longer met or that sudden barriers prevent access to a reward (Psyrdellis et al., 2016), were video recorded for further behavioural analyses (Fig. 1). The behaviours indicating 'frustration' and 'attentive state' (see description in Tables 3,4,5) were scored using the software The Observer XT v.11 (Noldus) as "State Events", by an observer who was blind to the groups (crib-biters or controls). For each behaviour, its duration relative to the total recording time was calculated. Frustration and attentive states are known to be indicators of stress (Valenchon et al., 2013; Psyrdellis et al., 2016).

2.3.3. *Physiological measures*

We assessed the activity of the hypothalamic-pituitary-adrenal stress pathway (HPA) by collecting salivary cortisol at three time points during the experiment (Fig. 1). Saliva was collected with Salivette cotton rolls placed loosely onto the tongue of the horse for one minute using forceps. At the end of the test, the Salivettes were centrifuged for six minutes at 5000 rpm with a Hettich EBA 20 centrifuge. Samples were stored at −20 °C before being sent for analyses (Salimetrics, USA). Concentrations of cortisol were determined with the Salimetrics® High Sensitivity Salivary Cortisol Enzyme Immunoassay kit, a direct enzyme immunoassay without extraction that has been validated for equine saliva (Schmidt et al., 2010).

Concerning the sympathomedullary stress pathway (SAM) axis, we assessed two physiological measures, the heart rate (HR) and the root mean square of successive inter-beat interval differences (RMSSD) (von Borell et al., 2007). The ECG trace was collected continuously using a wireless, non-invasive monitor (MLE120X Bioharness Telemetry System, Zephyr) fixed to a surcingle placed around the horse's heart girth,

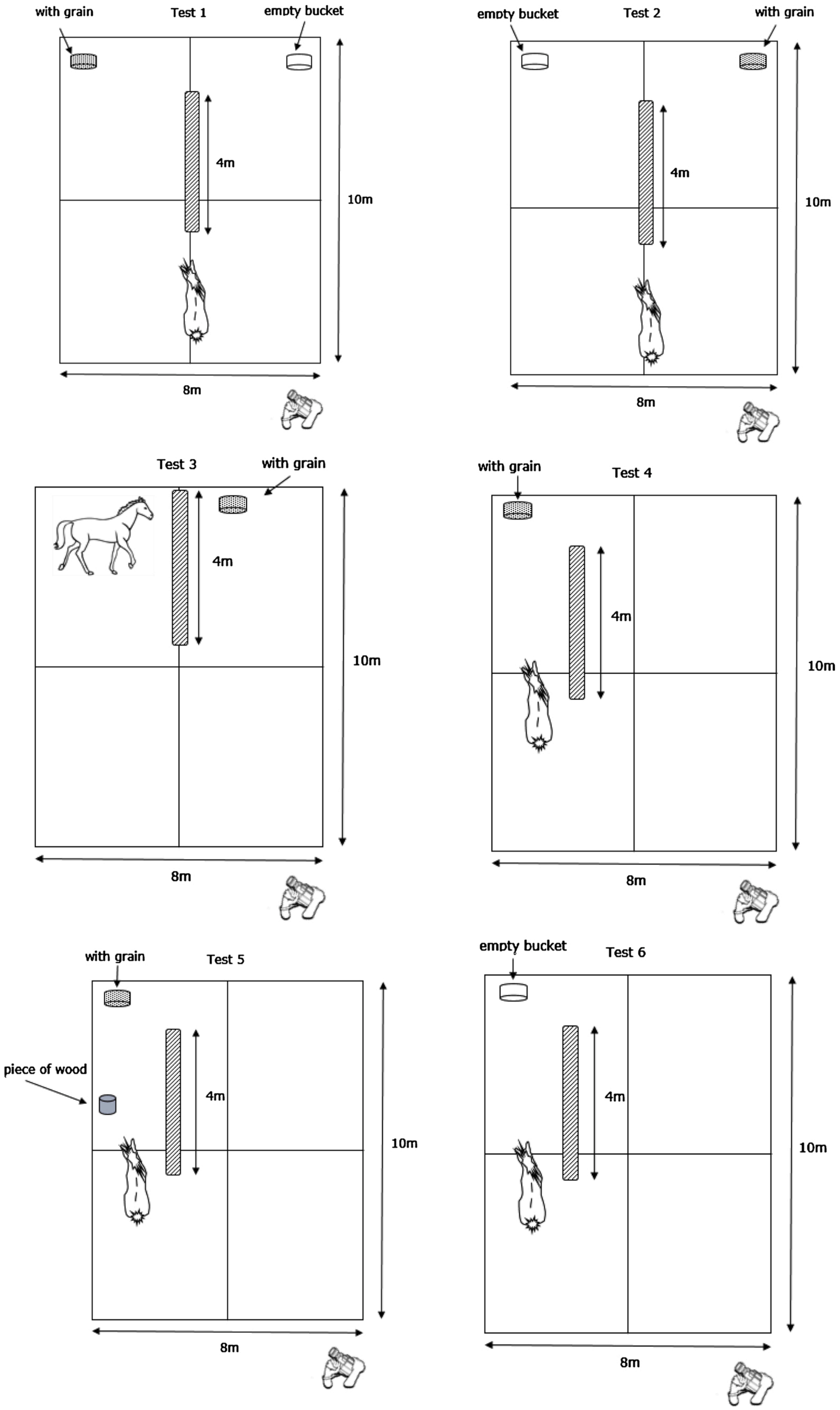


Fig. 2. Schematic representations of the six tests. The camera was positioned at the bottom right. The test arena was surrounded by solid fences (large rectangle). During *Tests 1* and *2*, subjects were released in front of the fence. In *Test 3*, subjects were released one metre perpendicular to the fence and at the start of the corridor in *Tests 4-6*.

Table 2
Crib-biting events per horse in the different learning tests. The number of the horses corresponds to those shown in Table 1. The two last rows indicate the sum of the crib-biting events per test (“CBnb”) for all the crib-biters, and the number of horses per test that crib-bit (“Nb”).

Crib-biters	Test1	Test2	Test 3	Test 4	Test 5	Test 6
1	0	0	0	0	0	0
2	0	0	0	0	0	1
3	1	0	0	0	0	8
4	0	0	128	3	4	53
5	0	0	0	0	0	0
6	12	0	1	0	0	2
7	0	0	11	0	0	11
8	0	2	0	0	0	0
9	0	2	0	0	0	0
10	0	0	0	0	0	7
11	0	2	6	0	0	7
12	0	0	0	0	0	0
13	0	0	0	0	0	0
14	2	1	0	0	0	2
15	0	0	0	0	0	2
16	9	11	0	0	0	1
17	0	0	0	0	0	30
18	0	0	7	0	1	2
19	0	0	0	0	0	0
CBnb	24	18	153	3	5	126
Nb	4	5	5	1	2	12

which produces more reliable heart-rate (HR) measures compared to alternative methods (Parker et al., 2010). The data were transmitted in real time during the tests to a laptop using LabChart software v.7.2 (ADInstrument) for later analyses. This allowed us to enter comments into the software during the experiment to indicate when each learning test started (Fig. 1). We then extracted HR and RMSSD from good-quality sections with clearly visible heartbeats on the ECG trace. Due to the different duration of the tests, when possible, we selected five short selections of about 10 s each for *Tests 1–3* and *6*, three periods of 10 s for *Test 4* and one period of 10 s for *Test 5* (Reefmann et al., 2009; Briefer et al., 2015). These short selections allowed us to identify even short-term changes in physiology linked to each test, as these measures rapidly return to baseline after transient perturbations (von Borell et al., 2007). We checked visually that the software tracked the heartbeats properly and extracted HR and inter-heartbeat (RR) intervals (in ms). RR intervals were then used to calculate RMSSD (ms). If atrioventricular blocks were observed in the signal, we excluded the respective sections of the ECG trace (Schmidt et al., 2010). For one crib-biter and one control, we were unable to obtain HR and RMSSD data due to the bad quality of the signal.

2.4. Statistical analysis

2.4.1. Learning progress

To assess if horses improved their performance throughout the different trials in the learning tests (*Test 1−4* and *6*), we compared the time taken by the horses to reach the bucket filled with grain (*Test 1−4*), or the time they spent waiting in front of the corridor (*Test 6*), during the first and the last trial of each test (“Time”), using linear mixed-effects models (LMM; lmer function, lme4 library). These LMMs included “Time” as a response variable. The fixed and control factors are listed and described in Tables 3 to 4 . To control for repeated measurements of the same subjects, the identity of the horses nested within the farms where they were housed was included as a random factor. We also investigated if the horses went less to look into the empty bucket over

Table 3
Abbreviations of the behavioural measures collected during the different tests and of the other parameters and factors used in the different analyses.

Abbreviations	Definition	Test
Behaviours scored through direct observation		
“Time” (s)	Time to go to the bucket filled with grain	*Tests 1−5*
	Time spent waiting in front of the corridor	*Test 6*
“CBnb” (n)	Number of crib-biting events per test	*Tests 1−6*
“Side” (b)	Side of the fence the horse takes to go to the bucket, i.e. left or right	*Tests 1−2*
“Bucket” (b)	Looking or not into the empty bucket	*Tests 1−2, 6*
“Touch” (b)	Touching or not the piece of wood	*Test 5*
“Wood” (b)	Looking or not at the piece of wood	*Test 5*
Behaviours scored from the videos		
“Frustr” (p)	Pushing the solid fence with the nose or the body, turning the empty bucket over with the nose or foreleg, shaking the head	*Test 1*
	Scratching the ground with the foreleg, pushing the solid fence with the nose or the body, turning the empty bucket over with the nose or foreleg	*Test 2*
	Scratching the ground with the foreleg, pushing the fence with the nose or the body	*Test 3*
	Scratching the ground with the foreleg, turning the empty bucket over with the leg or the nose	*Test 6*
“Att” (p)	Standing attentive, without moving, with the head and neck up and the ears pointing forward	*Tests 1−3, 6*
Other parameters		
Group (“CB” or “C”) (c)	Crib-biting and control group	*Tests 1−6*
Group (“A”, “B”, “C”) (c)	A = Crib-biting individuals that showed crib-biting during a test; B = Crib-biting individuals that did not show crib-biting during a test	*Tests 1−6*
“Trial” (n)	Trial number for each test. 1–8 for *Tests 1, 2* and *6*; 1–5, for *Test 3*; 1–3, for *Test 4*	*Tests 1−4, 6*
“Sideb” (b)	Side of the bucket that contained food, i.e. left or right	*Tests 1−2*
“HR”(a)	Heart rate, average value per test	*Tests 1−6*
“RMSSD” (a)	Root mean square of successive inter-beat interval differences, average value per test	*Tests 1−6*
“Part” (c)	Part 1	*Tests 1−3*
	Part 2	*Tests 4−6*
“Testnb” (n)	Test order: 1–7 for Part 1, including personality tests (1−4) analysed in another study (Briefer Freymond et al., 2019)	*Test 1−3*
	Test order: 8–12 for Part 2, including personality tests (11−12) analysed in another study (Briefer Freymond et al., 2019)	*Test 4−6*
“Hour” (n)	Hour at which a saliva sample was collected for cortisol analyse	
“Base” (n)	Cortisol value measured from the first saliva sample collected before the tests, i.e. baseline for cortisol analyses	
p: proportion; n: number; s : second, b: binary; c: category; a: average		

Table 4
Fixed and control factors included in the models used to analyse the behaviours scored through direct observation and the learning progess. The abbreviations are described in Table 3. The crosses indicate which factors or interaction between factors ('*') were used in the models to investigate the variation observed in the behavioural measures. The response variable was Time for *Test 1-6*, Wood and Touch for *Test 5* and Bucket for *Test 2* and *6*. In this Table, Group is used for Group CB-C or. ABC.

Factors		Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
Fixed	Group	x	x	x	x	x	x
	Trial	x	x	x	x		x
	Trial * Group	x	x	x	x		x
	Bucket * Group	x	x				x
	Trial * Bucket	x	x				x
Control	CBnb	x	x	x	x	x	x
	Age	x	x	x	x	x	x
	Sex	x	x	x	x	x	x
	Side	x	x				
	Arena	x	x	x	x	x	x
	Sideb	x	x				
	Bucket	x	x				x

Table 5
Fixed and control factors included in the models used to analyse the physiological measures. The abbreviations are described in Table 3. The crosses indicate which factors or interactions between factors ('*') were used in the models investigating the variation observed in the physiological parameters (response variable: Cort, HR, RMSSD). In this Table, Group is used for Group CB-C or. ABC.

Factors		Cort	HR	RMSSD
Fixed	Group	x	x	x
	Part * Group	x		
	Testnb * Group		x	x
Control	Part	x		
	Base	x		
	Sex	x	x	x
	Age	x	x	x
	Testnb		x	x
	Hour	x		
	CBnb		x	x

trials for *Test 2* and *Test 6* (see 2.4.2).

2.4.2. Differences between crib-biters and controls

2.4.2.1. Behaviours scored through direct observation. To compare the behaviour of crib-biters ("CB") and controls ("C") during the learning tests, we investigated group differences in the time needed to reach the bucket (*Tests 1–4*) or in the time spent waiting in front of the corridor (*Test 6*) ("Time") using linear mixed-effects models (LMM; lmer function) and generalised linear mixed models (GLMM; glmer function; lme4 library; multcomp library (Pinheiro and Bates, 2000) in R 3.0.2 (Table 4)). We also investigated group differences in looking or not at the empty bucket ("Bucket") (*Tests 2, 6*), in looking or not at the piece of wood ("Wood") and in touching or not the piece of wood ("Touch") (*Test 5*) (binary responses) using GLMMs (Tables 3 and 4). The LMMs included "Time" as a response variable (*Test 1, 2, 3, 5*). The GLMMs included "Time" transformed as a binary variable (*Tests* 4), "Bucket" (*Tests 2, 6*), "Wood" or "Touch" (*Test 5*) as a response variable (one model per variable) (Tables 3 and 4). The fixed, control factors and interaction terms included in all models are listed and described in Tables 3,4. To control for repeated measurements of the same subjects, the identity of the horses nested within the farms where they were housed was included as a random factor. For *Test 5*, only Farms were included as a random factor, as there was only one value for each horse. All parameter abbreviations are defined and summarised in Table 3.

Over all six learning tests, 14 of 19 crib-biters displayed stereotypic behaviour at least once (Tables 1 and 2). We therefore then reran the LMMs and GLMMs described above to compare, this time, crib-biters that did ("A") and did not ("B") show stereotypic behaviour in each test (Tables 1 and 2) and control subjects ("C").

2.4.2.2. Behaviours scored from the videos. Two response variables were analysed (*Test 1, 2, 3 and 6*); frustration ("Frustr") and attentive state ("Att"), using either LMM or GLMM depending on the transformation needed (Tables 3 and 4). The group (Group CB-C or Group ABC; Tables 3 and 4) was considered as a fixed factor. The other control parameters are listed and defined in Tables 3 and 4. Farm was added as a random factor.

2.4.2.3. Physiological measures. To compare the physiology of crib-biters and controls during the learning tests, we used LMMs, including as a response variable (3 separate models) the actual saliva cortisol value ("Cort"), or the average HR or RMSSD value for each learning test (hereafter "HR" and "RMSSD", respectively; Tables 3 and 5). The group (Group CB-C or Group ABC; Tables 3 and 5) was considered as a fixed parameter. For the cortisol analyses, horses were included in Group A (crib-biters that did crib-bite) when they did crib-bite in the corresponding part of the experiment where the cortisol was measured. For HR and RMSSD, because we calculated an average HR and RMSSD value for each test, crib-biters were included in Group A as soon as they started to crib-bite in the corresponding part. The fixed, control factors and interaction terms are listed in Tables 3 and 5. To control for repeated measurements of the same subjects, the identity of the horses nested within their respective farms was included as a random factor. In addition, for HR and RMSSD, the part (Part 1 or 2; Table3) was included as a crossed random factor.

For all models, the residuals were checked graphically for normal distribution and homoscedasticity. To satisfy model assumptions, we used a log transformation for "Time" in *Test 6,* for "RMSSD", and for "Att" in *Test 1*, and a square-root transformation for "Att" in *Test* 2 and *Test* 6 and for "Frustr" in *Test 6* (see Table 3 for abbreviations). "HR" and "Time" in *Test 2* were cube-root transformed. Finally, "Time" in *Test 1* and *3,* which did not meet statistical assumption despite log, square-root or cube-root transformation, was split in six categories of equal duration (i.e., 1: 1–30 s, 2: 31–60 s, 3: 61–90 s, 4: 91–120 s, 5: 121–150 s, 6: 151–180 s). All the resulting parameters satisfying model assumptions were then input into linear mixed-effects models (lmer function). Some parameters did not meet the statistical assumptions despite transformation. They were thus transformed to binomial data as follows; value equal or higher than median = 1, and value lower than median = 0. This was the case for "Time" in *Test 4,* for "Frustr" in *Test* 2 and 3 and for "Att" in *Test* 3. The parameters scored as binomial (Bucket, Wood, Touch), as well as parameters transformed to binomial data, were input into GLMM fit with binomial family distribution and logit link function (glmer function).

For both the LMMs and GLMMs, we used a standard model simplification procedure by removing each non-significant term, until the deletion caused a reduction in goodness of fit (at which point, the term was left in the model). *P*-values were calculated based on Satterthwaite's approximations (anova function, lmerTest package in R). The significance level of the factors was set at $\alpha = 0.05$ for all models.

For all the models, when significant interaction effects between fixed and/or control factors were found and when Group ABC had a significant effect, further post-hoc analysis were carried out using LMMs and GLMMs. In these cases, a Bonferroni correction was applied. Only the result of the fixed factors of interest (and not of the control factors) are mentioned in the results.

3. Results

One control horse was not motivated by rewards and did not perform any of the learning tests. It was therefore removed from the analyses. The sample size is thus eventually of $N = 19$ crib-biters and 17 control.

3.1. Learning progress

For each test, there was a significant difference between the first and the last trial of the test in the time required to go to the bucket (*Test 1−4*), or in the time spent waiting in front of the corridor (*Test 6*) (“Time”; LMMs: effect of Trial on Time; $p \leq 0.04$ for all; for *Test 1−4,* slope estimate range: –8.32 to -0.16; for *Test 6:* slope estimate ± SD = 0.40 ± 0.03). Therefore, all horses improved their performances throughout the trials as shown by a decrease in time between the first and the last trial for *Tests 1−4,* and an increase in the time spent waiting in front of the fence for *Test 6.* In both *Test 1* and *Test 6,* for all the horses, their likeliness to go look at the empty bucket (“Bucket”) decreased with Trial (GLMM: effect of Bucket on Trial, $X^2_1 = 48.11$, $p < 0.0001$; slope estimate ± SD: -0.53 ± 0.09, in *Test 1* and GLMM: effect of Bucket on Trial, $X^2_1 = 62.44$, $p < 0.0001$; slope estimate ± SD: -0.49 ± 0.07, in *Test 6*), indicating that horses improved their performance over the trials, as shown by a decrease in their propensity to go look at the empty bucket.

3.2. Differences between crib-biters and controls

3.2.1. Behaviours scored through direct observation

In the *place learning* test (*Test 1*), there was no difference between crib-biters (“CB”) and controls (“C) in the time taken to go to the bucket containing grains (“Time”) (LMM: effect of Group CB-C on Time, $X^2_1 = 0.74$, $p = 0.39$). However, there was an effect on Time of the interaction between Group ABC (“A”, crib-biters that did crib-bite in Part 1; “B”, crib-biters that did not crib-bite in Part 1; and “C”, controls; Table 2) and the corresponding trial (“Trial”; LMM: interaction effect between Groups ABC and Trial on Time, $X^2_2 = 13.45$, $p = 0.001$), as well as of the interaction between Group ABC and if the horses went to look at the empty bucket or not (“Bucket”) (LMM: interaction effect between Group ABC and Bucket on Time, $X^2_2 = 12.68$, $p = 0.002$). However, post-hoc analyses per Trial and Bucket did not show further significant differences between groups (LMMs: $p = 1.00$ for all).

In the *reversal learning* test (*Test 2*), there was no effect of Group CB-C, and of Group ABC, on Time (LMM: effect of Group CB-C on Time, $X^2_1 = 1.36$, $p = 0.24$; LMM: effect of Group ABC on Time, $X^2_2 = 1.52$, $p = 0.47$). Because the side where the bucket containing food and the empty bucket was situated was exchanged between Test 1 and Test 2 (Fig. 2), we also analysed if there was a difference between groups in Bucket. This effect was not significant neither for Group CB-C, nor for Group ABC (GLMM: effect of Group CB-C on Bucket, $X^2_1 = 0.0012$, $p = 0.97$; effect of Group ABC on Bucket, $X^2_2 = 1.06$, $p = 0.59$).

In the *detour learning* test (*Test 3*), there was an effect on Time of the interaction between Group CB-C and Trial (LMM: interaction effect between Group CB-C and Trial on Time, $X^2_1 = 4.05$, $p = 0.04$), as well as of the interaction between Group ABC and Trial (LMM: interaction effect between Group ABC and Trial on Time, $X^2_2 = 13.45$, $p = 0.001$). However, further post-hoc analyses per Trial did not show any significant effect of the factor Group on Time, and post-hoc analyses per Group did not show any significant effect of Trial on Time (LMMs: $p \geq 0.31$ for all).

In the *food motivation* test (*Test 4*), there was an effect on Time of the interaction between Group CB-C and Trial (GLMM: interaction effect between Group CB-C and Trial on Time, $X^2_1 = 6.96$, $p = 0.008$), as well as of the interaction between Group ABC and Trial (GLMM: interaction effect between Group ABC and Trial on Time, $X^2_2 = 7.15$, $p = 0.03$). However, further post-hoc analyses per Trial did not show any significant effect of the factor Group on Time, and post-hoc analyses per Group did not reveal any significant effect of Trial on Time (GLMMs: $p \geq 0.07$ for all).

In the *distractibility* test (*Test 5*), there was no effect of Group CB-C, and of Group ABC, on either looking at the piece of wood or not (“Wood”) or on touching the piece of wood or not (“Touch”) (GLMM: effect of Group CB-C on Wood, $X^2_1 = 46.461$, $p = 0.08$; effect of Group CB-C on Touch, $X^2_1 = 0.07$, $p = 0.80$; effect of Group ABC on Wood, $X^2_2 = 46.03$, $p = 0.13$; effect of Group ABC on Touch, $X^2_2 = 0.43$, $p = 0.80$).

In the *extinction* test (*Test 6),* there was no effect of Group CB-C, and of Group ABC, on Time (LMM: effect of Group CB-C on Time, $X^2_1 = 0.22$, $p = 0.64$; effect of Group ABC on Time, $X^2_2 = 3.37$ $p = 0.19$). Because in this test, the bucket previously containing food had been replaced by an empty bucket, we also analysed if there was a difference between groups in Bucket. There was no effect of Group CB-C, nor of Group ABC on Bucket (GLMM: effect of Group CB-C on Bucket, $X^2_1 = 0.36$, $p = 0.55$; effect of Group ABC on Bucket, $X^2_2 = 2.27$, $p = 0.32$).

3.2.2. Behaviours scored from the videos

In the *place learning* test *(Test 1),* there was no effect of either Group CB-C or Group ABC on frustration behaviours (“Frustr”) (GLMM: effect of the Group CB-C on Frustr, $X^2_1 = 1.68$, $p = 0.19$; effect of the Group ABC on Frustr, $X^2_2 = 2.12$, $p = 0.35$). There was also no effect of either Group CB-C or Group ABC on attentive state “Att” (LMM: effect of Group CB-C on Att, $X^2_1 = 0.001$, $p = 0.98$; effect of Group ABC on Att, $X^2_2 = 2.29$, $p = 0.32$).

In the *reversal learning* test *(Test 2),* there was no effect of Group CB-C on Frustr (GLMM: effect of the Group CB-C on Frustr, $X^2_1 = 0.09$, $p = 0.77$). However, there was an effect of Group ABC on this behaviour (GLMM: effect of the Group ABC on Frustr, $X^2_2 = 7.40$, $p = 0.02$). Further post-hoc comparisons showed that more horses from Group B did show frustration behaviours in this test compared to Group A horses (mean proportion of horses showing frustration behaviour ± SD: Group A, 0.40 ± 0.52; Group B, 0.78 ± 0.44; GLMM: effect of Group A versus B on Frustr, $X^2_1 = 7.76$, $p = 0.015$), while there was no difference between horses from Group A and C nor between horses from Group B and C in this behaviour (mean proportion of horses showing frustration behaviour ± SD: Group C, 0.50 ± 0.51; GLMM: effect of Group A versus C on Frustr, $X^2_1 = 0.18$, $p = 1.0$; effect of Group B versus C on Frustr, $X^2_1 = 4.33$, $p = 0.12$). Finally, there was no effect of either Group CB-C or Group ABC on Att (LMM: effect of Group CB-C on Att, $X^2_1 = 1.58$, $p = 0.21$; effect of Group ABC on attentive state, $X^2_2 = 3.33$, $p = 0.19$).

In the *detour learning* test *(Test 3)*, there was no effect of either Group CB-C or Group ABC on Frustr (GLMM: effect of Group CB-C on Frustr, $X^2_1 = 0.091$, $p = 0.76$; effect of Group ABC on Frustr, $X^2_2 = 0.39$, $p = 0.82$). There was also no effect of either Group CB-C or Group ABC on Att (GLMM: effect of Group CB-C on Att, $X^2_1 = 0.07$, $p = 0.79$; effect of Group ABC on Att, $X^2_2 = 0.50$, $p = 0.78$).

In the *extinction* test (*Test 6*), there was no effect of either Group CB-C or Group ABC on Frustr (LMM: effect of Group CB-C on Frustr, $X^2_1 = 2.60$, $p = 0.11$; effect of Group ABC on Frustr, $X^2_2 = 4.83$, $p = 0.09$). There was also no effect of either Group CB-C or Group ABC on Att (LMM: effect of the Group CB-C on Att, $X^2_1 = 0.23$, $p = 0.63$; effect of the Group ABC on Att, $X^2_2 = 2.75$, $p = 0.25$).

3.2.3. Physiological measures

3.2.3.1. Cortisol. There was no difference in actual cortisol levels (“Cort”) between groups CB and C (LMM: effect of Groups CB-C on cortisol, $X^2_1 = 0.03$, $p = 0.87$). However, there was an interaction effect between the experimental part (Part 1 or 2; “Part”) and Group ABC on Cort (LMM: interaction effect between Part and Group ABC on Cort, $X^2_2 = 8.56$, $p = 0.014$). Post-hoc comparisons showed that Cort differed between Groups ABC after Part 1 (LMM: effect of Group ABC on Cort in Part 1, $X^2_2 = 8.71$, $p = 0.03$), but not after Part 2 (LMM: effect of Group ABC on Cort in Part 2, $X^2_2 = 0.90$, $p = 1.00$; Fig. 3). Further two-by-two comparisons showed that during Part 1, Group A had a lower actual

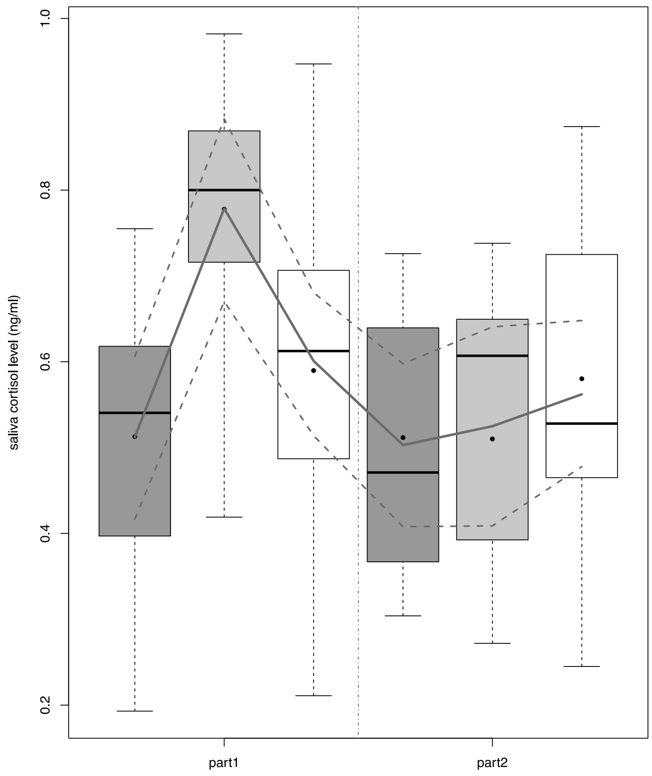


Fig. 3. Saliva cortisol values as a function of the crib-biting group. Saliva cortisol level in crib-biters that did crib-bite (dark grey; $N = 10$ in Part 1 and $N = 12$ in Part 2), crib-biters that did not crib-bite (light grey; $N = 9$ in Part 1 and $N = 7$ in Part 2) and control horses (white; $N = 18$); box-and-whiskers plot (the horizontal line shows the median, the box extends from the lower to the upper quartile, and the whiskers to 1.5 * the interquartile range above the upper quartile or below the lower quartile). The black dots indicate the means. Significant differences between Phase are indicated as * $p \leq 0.05$. The lines show the model estimates (middle line) and 95 % confidence intervals (upper and lower dotted line).

cortisol value than Group B (effect of Group A versus B on Cort; LMM: $X^2_2 = 11.012$, $p = 0.024$), and than Group C (effect of Group A versus C on Cort; LMM: $X^2_2 = 12.55$, $p = 0.012$). By contrast, Group B and Group C did not differ (effect of Group B versus C on Cort; $X^2_2 = 5.75$, $p = 0.17$; Fig. 3).

3.2.3.2. Heart rate (HR). There was no effect of Group CB-C, nor of Group ABC, on the average heart-rate ("HR") value per test (LMM: effect of Groups CB-C on HR, $X^2_1 = 0.501$, $p = 0.48$; effect of Groups ABC on HR, $X^2_2 = 2.05$, $p = 0.36$).

3.2.3.3. Root mean square of successive inter-beat interval differences (RMSSD). There was no effect of Group CB-C, nor of Group ABC, on the average value of RMSSD per test (LMM: effect of Group CB-C on RMSSD, $X^2_1 = 0.16$, $p = 0.69$; effect of Groups ABC on HR, $X^2_2 = 1.72$, $p = 0.42$).

For all the models, the effect of the fixed factors of interest and interactions not mentioned above were not significant.

4. Discussion

In the current study, we investigated the learning abilities of a large sample of horses, half of which showed a stereotypy (crib-biting behaviour). The horses performed six spatial learning tasks, including some that have been considered as challenging for stereotypic animals, that is, reversal learning, detour and extinction (McBride et al., 2017). Across experiments, all subjects (crib-biters and controls) improved their performance, with no between-group differences in learning. However, within the stereotypic group, we found differences between subjects that did or did not perform stereotypic behaviour. In particular, a smaller proportion of crib-biters that did crib-bite showed behaviours indicating frustration compared to those that did not crib-bite in the reversal learning test. Accordingly, these horses had lower cortisol values after the first experimental part compared to those that did not crib-bite and controls. These results corroborate previous findings (Briefer Freymond et al., 2015), which suggest that stereotypic behaviour is an adaptive strategy that enables a reduction of stress levels.

Our results did not show any difference in performance between stereotypic and non-stereotypic horses. One explanation could be that the learning tests we used (spatial learning tests) were not challenging enough to result in any learning differences between our two groups. Indeed, horses have been shown to excel in finding food sources using spatial cues, suggesting that this might be an easy task for them (Martin et al., 2006; Hothersall et al., 2010; Brubaker and Udell, 2016). Second, in previous studies suggesting an effect of habit formation, horses had been trained for longer (e.g. 40 trials), because the task was more difficult for them to learn, compared to our study (Parker et al., 2008). Such habit formation might have not occured within the limited number of trials used in our study.

Another explanation for the absence of group differences in performance in our study could be that the stress levels of stereotypic subjects might have decreased resulting in a lower dependency on rewards after executing stereotypic behaviour on the solid fence during the tests. This could have, as a result, increased stereotypic subjects' performance. Indeed, there is a general consensus that, across species, cognitive performances decrease with increasing levels of stress (Valenchon et al., 2013). Similarly, learning under stress can negatively impact on subsequent performance (in humans: Schwabe and Wolf, 2010), possibly because stress acts as a distractor during encoding and, instead, favours inflexible habit formation (Yin and Knowlton, 2006; McBride et al., 2017). These results corroborate our previous findings (Briefer Freymond et al., 2018), which suggest that allowing crib-biters to crib-bite may improve their learning performance through a decrease in stress, even in highly challenging tasks such as reversal learning based on visual cues.

The lower stress levels experienced by crib-biters in our study was confirmed by our physiological analyses. Indeed, although our results did not reveal any difference in cortisol between crib-biters and control horses, crib-biters that did crib-bite had lower cortisol levels after the first experimental part, indicating lower stress levels, than crib-biters that did not crib-bite and even control horses. This suggests that crib-biting functions as a coping strategy that helps horses to deal with stressful (e.g. frustrating) situations (McBride and Cudderford, 2001; Briefer Freymond et al., 2015) and to avoid further pathogenic outcomes (Cabib and Puglisi-Allegra, 2012). This suggestion is strengthened by the fact that a smaller proportion of crib-biters that did crib-bite showed behaviours indicating frustration than crib-biters that did not crib-bite in the reversal learning test. Stereotypic behaviour may therefore act as a way to reduce stress and frustration and hence increase well-being (Mason and Latham, 2004; Briefer Freymond et al., 2015). We suggest that the absence of cortisol difference between crib-biters that did crib-bite, crib-biters that did not crib-bite and controls after the second experimental part could be due to the fact that the learning tests (*Tests 4–6*) were followed by other tests aimed at investigating group differences in personality (Briefer Freymond et al., 2019), which likely induced less frustration compared with the learning tests.

Why some crib-biters in our study refrained from showing stereotypic behaviour during frustrating situations is not clear. One possibility is that these horses were at an earlier stage of stereotypy development and, correspondingly, might not have had suffered from a clear dopaminergic dysfunction in some parts of the *basal ganglia* yet. Further experimental studies are required to test this hypothesis, by exploring the impact of preventing stereotypic behaviour to different degrees on stress parameters. Such parameters could include, for instance, spontaneous blink rate (Roberts et al., 2015), an indicator of dopamine transmission and striatal functioning as well as of dopaminergic component of sustained attention and fatigue in humans (Maffei and Angrilli, 2018).

5. Conclusion

In this study, we did not find clear differences in learning capacities between crib-biting and control horses during a battery of learning tests. Our tests all involved spatial problems to solve, with limited training, which might be more ecologically relevant for horses, and thus easier, compared to paradigms used in previous studies (Hemmings et al., 2007; Parker et al., 2008, 2009). In our study, all crib-biters had the possibility to execute stereotypic behaviour during testing. Overall, we could suggest that executing stereotypic behaviour lowered both stress (as assessed by cortisol levels) and frustration, indicating that engaging in stereotypic behaviour is a strategy to prevent stress, which might then help crib-biters to better solve learning tasks. Our results therefore lead us to recommend that, as previously suggested (Hausberger et al., 2007; Briefer Freymond et al., 2018), stereotypic horses, and perhaps also human showing stereotypic behaviours, may require specific training adapted to their needs. Letting stereotypic individuals the possibility to perform their stereotypy might be one specific feature to incorporate in learning protocols, which might then allow these individuals to perform successfully.

Author’s contribution statement

S.B.F. and S.B. carried out the experiment. A.R. scored the videos. S.B.F. wrote the manuscript and performed the statistical analyses. E.F.B., K.Z and I.B. contributed to the study design and edited the manuscript. E.F.B supervised the project. All authors read and approved the final manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgments

We are grateful to the owners of the horses, who offered to participate in this study. We are also grateful to Matthew Parker, who helped us with the design of the experiment.

References

Albright, J., Mohammed, H., Heleski, C., Wickens, C., Houpt, K., 2009. Crib-biting in US horses: breed predispositions and owner perceptions of aetiology. Equine Vet. J. 41, 455–458.

Bhatnagar, S., Vining, C., 2003. Facilitation of hypothalamic-pituitary-adrenal responses to novel stress following repeated social stress using the resident/intruder paradigm. Horm. Behav. 43, 158–165.

Boyd, B.A., McBee, M., Holtzclaw, T., Baranek, G.T., Bodfish, J.W., 2009. Relationships among repetitive behaviors, sensory features, and executive functions in high functioning autism. Res. Autism Spectr. Disord. 3, 959–966.

Briefer, E.F., Maigrot, A.-L., Mandel, R., Freymond, S.B., Bachmann, I., Hillmann, E., 2015. Segregation of information about emotional arousal and valence in horse whinnies. Sci. Rep. 4, 9989.

Briefer Freymond, S., Bardou, D., Briefer, E.F., Bruckmaier, R., Fouché, N., Fleury, J., Maigrot, A.L., Ramseyer, A., Zuberbühler, K., Bachmann, I., 2015. The physiological consequences of crib-biting in horses in response to an ACTH challenge test. Physiol. Behav. 151, 121–128.

Briefer Freymond, S., Ruet, A., Grivaz, M., Fuentes, C., Zuberbühler, K., Bachmann, I., Briefer, E.F., 2018. Stereotypic horses (Equus caballus) are not cognitively impaired. Anim. Cogn.

Briefer Freymond, S., Bardou, D., Beuret, S., Zuberbuhler, K., Bachmann, I., Briefer, E.F., 2019. Elevated sensitivity to tactile stimuli in stereotypic horses. Front. Vet. Sci. 6, 162.

Brubaker, L., Udell, M.A.R., 2016. Cognition and learning in horses (Equus caballus): what we know and why we should ask more. Behav. Processes 126, 121–131.

Cabib, S., Puglisi-Allegra, S., 2012. The mesoaccumbens dopamine in coping with stress. Neurosci. Biobehav. Rev. 36, 79–89.

Cabib, S., Giardino, L., Calzà, L., Zanni, M., Mele, A., Puglisi-Allegra, S., 1998. Stress promotes major changes in dopamine receptor densities within the mesoaccumbens and nigrostriatal systems. Neuroscience 84, 193–200.

Cooper, J., Mason, G., 1998. The identification of abnormal behaviour and behavioural problems in stabled horses and their relationship to horse welfare: a comparative review. Equine Vet. J. 30, 5–9.

Dallaire, J.A., Meagher, R.K., Diez-Leon, M., Garner, J.P., Mason, G.J., 2011. Recurrent perseveration correlates with abnormal repetitive locomotion in adult mink but is not reduced by environmental enrichment. Behav. Brain Res. 224, 213–222.

Frith, C.D., Done, D.J., 1983. Stereotyped responding by schizophrenic patients on a two-choice guessing task. Psychol. Med. 13, 779–786.

Garner, J.P., Mason, G.J., 2002. Evidence for a relationship between cage stereotypies and behavioural disinhibition in laboratory rodents. Behav. Brain Res. 136, 83–92.

Garner, J.P., Meehan, C.L., Mench, J.A., 2003. Stereotypies in caged parrots, schizophrenia and autism: evidence for a common mechanism. Behav. Brain Res. 145, 125–134.

Giuliana, V.G., Roberto, G., 1986. Evidence of heredity of cribbing, weawing and stall-walking in thoroughbred horses. Livest. Prod. Sci. 14, 91–95.

Hausberger, M., Gautier, E., Müller, C., Jego, P., 2007. Lower learning abilities in stereotypic horses. Appl. Anim. Behav. Sci. 107, 299–306.

Hemmings, A., McBride, S.D., Hale, C.E., 2007. Perseverative responding and the aetiology of equine oral stereotypy. Appl. Anim. Behav. Sci. 104, 143–150.

Hemmings, A., Parker, M.O., Hale, C., McBride, S.D., 2018. Causal and functional interpretation of mu- and delta-opioid receptor profiles in mesoaccumbens and nigrostriatal pathways of an oral stereotypy phenotype. Behav. Brain Res. 353, 108–113.

Hothersall, B., Gale, E., Harris, P., Nicol, C., 2010. Cue use by foals (Equus caballus) in a discrimination learning task. Anim. Cogn. 13, 63–74.

Ijichi, C.L., Collins, L.M., Elwood, R.W., 2013. Evidence for the role of personality in stereotypy predisposition. Anim. Behav. 85, 1145–1151.

Izquierdo, A., Brigman, J.L., Radke, A.K., Rudebeck, P.H., Holmes, A., 2017. The neural basis of reversal learning: an updated perspective. Neuroscience 345, 12–26.

Jentsch, J.D., Roth, R.H., Taylor, J.R., 2000. Object retrieval/detour deficits in monkeys produced by prior subchronic phencyclidine administration: evidence for cognitive impulsivity. Biol. Psychiatry 48, 415–424.

Kabadayi, C., Bobrowicz, K., Osvath, M., 2017. The detour paradigm in animal cognition. Anim. Cogn. 1–15.

Koolhaas, J.M., 2011. Stress revisited: a critical evaluation of the stress concept. Neurosci. Biobehav. Rev. 35, 1291–1301.

Lesimple, C., Hausberger, M., 2014. How accurate are we at assessing others’ well-being? The example of welfare assessment in horses. Front. Psychol. 5, 21.

Lopez, B.R., Lincoln, A.J., Ozonoff, S., Lai, Z., 2005. Examining the relationship between executive functions and restricted, repetitive symptoms of Autistic Disorder. J. Autism Dev. Disord. 35, 445–460.

Luescher, U.A., McKeown, D.B., Halip, J., 1991. A cross-sectional study on compulsive behavior (stable vices) in horses. Equine Vet. J. 27, 14–18.

Maffei, A., Angrilli, A., 2018. Spontaneous eye blink rate: an index of dopaminergic component of sustained attention and fatigue. Int. J. Psychophysiol. 123, 58–63.

Martin, T.I., Zentall, T.R., Lawrence, L., 2006. Simple discrimination reversals in the domestic horse (Equus caballus): effect of discriminative stimulus modality on learning to learn. Appl. Anim. Behav. Sci. 101, 328–338.

Mason, G., Latham, N., 2004. Can’t stop, won’t stop: is stereotypy a reliable animal welfare indicator? Anim. Welf. 13, 57–69.

Mason, G., Rushen, J., 2006. Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare, second ed. CAB Interantional, Wallingford OXON.

McBride, S., Cudderford, D., 2001. The putative welfare-reducing effects of preventing equine stereotypic behaviour. Anim. Welf. 10, 173–189.

McBride, S.D., Hemmings, A., 2005. Altered mesoaccumbens and nigro-striatal dopamine physiology is associated with stereotypy development in a non-rodent species. Behav. Brain Res. 159, 113–118.

McBride, S.D., Parker, M.O., 2015. The disrupted basal ganglia and behavioural control: an integrative cross-domain perspective of spontaneous stereotypy. Behav. Brain Res. 276, 45–58.

McBride, S.D., Parker, M.O., Roberts, K., Hemmings, A., 2017. Applied neurophysiology of the horse; implications for training, husbandry and welfare. Appl. Anim. Behav. Sci. 190, 90–101.

McGreevy, P.D., Cripps, P.J., French, N.P., Green, L.E., Nicol, C.J., 1995. Management factors associated with stereotypic and redirected behaviour in the Thoroughbred horse. Equine Vet. J. 27, 86–91.

Muñoz, A., Torres, G., Briones, L., 2018. Prevalence of crib-biting and weaving in stabled horses tie stalls. Rev. Med. Vet. 65–72.

Parker, M., Redhead, E.S., Goodwin, D., McBride, S.D., 2008. Impaired instrumental choice in crib-biting horses (Equus caballus). Behav. Brain Res. 191, 137–140.

Parker, M., McBride, S.D., Redhead, E.S., Goodwin, D., 2009. Differential place and response learning in horses displaying an oral stereotypy. Behav. Brain Res. 200, 100–105.

Parker, M., Godwin, D., Eager, R.A., edhead, E.S., arlin, D.J., 2010. Comparison of Polar© heart rate interval data with simultaneously recorded ECG signals in horses. Comp. Exerc. Physiol. 6, 137–142.

Pinheiro, J.C., Bates, D.M., 2000. Mixed-Effects Models in S and S-PLUS. Springer-Verlag, New York, USA.

Psyrdellis, M., Pautassi, R.M., Mustaca, A., Justel, N., 2016. Cholinergic transmission underlies modulation of frustration by open field exposure. Pharmacol. Biochem. Behav. 140, 8–16.

Reefmann, N., Wechsler, B., Gygax, L., 2009. Behavioural and physiological assessment of positive and negative emotion in sheep. Anim. Behav. 78, 651–659.

Roberts, K., Hemmings, A., Moore-Colyer, M., Hale, C.E., 2015. Cognitive differences in horses performing locomotor versus oral stereotypic behaviour. Appl. Anim. Behav. Sci. 168, 37–44.

Saka, E., Goodrich, C., Harlan, P., Madras, B.K., Graybiel, A.M., 2004. Repetitive behaviors in monkeys are linked to specific striatal activation patterns. J. neurosci. Off. J. Soc. Neurosci. 24, 7557–7565.

Sambraus, H., 1985. Mouth-based anomalous syndromes. Ethology farm animals 391–422.

Schmidt, A., Möstl, E., Wehnert, C., Aurich, J., Müller, J., Aurich, C., 2010. Cortisol release and heart rate variability in horses during road transport. Horm. Behav. 57, 209–215.

Schwabe, L., Wolf, O.T., 2010. Learning under stress impairs memory formation. Neurobiol. Learn. Mem. 93, 183–188.

Schwabe, L., Wolf, O.T., 2011. Stress-induced modulation of instrumental behavior: from goal-directed to habitual control of action. Behav. Brain Res. 219, 321–328.

Turner, M., 1997. Towards an Executive Dysfunction Account of Repetitive Behaviour in Autism, Autism As an Executive Disorder. Oxford University Press, New York, NY, US, pp. 57–100.

Valenchon, M., Lévy, F., Fortin, M., Leterrier, C., Lansade, L., 2013. Stress and temperament affect working memory performance for disappearing food in horses, Equus caballus. Anim. Behav. 86, 1233–1240.

Vickery, S.S., Mason, G.J., 2005. Stereotypy and perseverative responding in caged bears: further data and analyses. Appl. Anim. Behav. Sci. 91, 247–260.

von Borell, E.H., 2001. The biology of stress and its application to livestock housing and transportation assessment. J. Anim. Sci. 79, 260–267.

von Borell, E., Langbein, J., Després, G., Hansen, S., Leterrier, C., Marchant-Forde, J., Marchant-Forde, R., Minero, M., Mohr, E., Prunier, A., Valance, D., Veissier, I., 2007. Heart rate variability as a measure of autonomic regulation of cardiac activity for assessing stress and welfare in farm animals A review, Physiology & Behavior. Stress Welfare Farm Animals 293–316.

Wolff, A., Hausberger, M., 1996. Learning and memorisation of two different tasks in horses: the effects of age, sex and sire. Appl. Anim. Behav. Sci. 46, 137–143.

Würbel, H., Stauffacher, M., 1994. Standard-Haltung für labormäuse–Probleme und lösungsansätze. Tierlaboratorium 17, 109–118.

Yin, H.H., Knowlton, B.J., 2006. The role of the basal ganglia in habit formation. Nat. Rev. Neurosci. 7, 464.